1-1-2005

Increased rates of molecular evolution in an equatorial plant clade: An effect of environment or phylogenetic nonindependence?

Jeremy M. Brown
The University of Texas at Austin

Gregory B. Pauly
The University of Texas at Austin

Recommended Citation

Brown, J., & Pauly, G. (2005). Increased rates of molecular evolution in an equatorial plant clade: An effect of environment or phylogenetic nonindependence?. *Evolution, 59* (1), 238-242. https://doi.org/10.1111/j.0014-3820.2005.tb00911.x

Evolution, 59(1), 2005, pp. 238–242

INCREASED RATES OF MOLECULAR EVOLUTION IN AN EQUATORIAL PLANT CLADE: AN EFFECT OF ENVIRONMENT OR PHYLOGENETIC NONINDEPENDENCE?

JEREMY M. BROWN[1,2] AND GREGORY B. PAULY[1,3]
[1]*Section of Integrative Biology, University of Texas, Austin, Texas, 78712*
[2]*E-mail: jembrown@mail.utexas.edu*
[3]*E-mail: gbpauly@mail.utexas.edu*

Abstract.—A recent study of environmental effects on rates of molecular evolution in the plant subgenus *Mearnsia* shows that species occurring in more equatorial latitudes have higher rates of substitution in rDNA sequences as compared to their more southerly congeners (Wright et al. 2003). However, we believe that the statistical approach employed by Wright et al. (2003) insufficiently accounts for the phylogenetic nonindependence of the species examined, given that all six equatorial species of *Mearnsia* form a clade. To distinguish between the effect of latitude and that of phylogenetic nonindependence, we have employed a variety of comparative approaches that use independent contrasts to test for an effect of environment across this entire subgenus. We find very little evidence for an effect of latitude on rate of molecular evolution using these approaches and believe that the shared evolutionary history of the clade is a plausible explanation of the apparent rate difference between equatorial and subequatorial *Mearnsia* species.

Key words.—Independent contrasts, latitudinal gradients, *Mearnsia*, penalized likelihood rate smoothing, relative rates test.

Received June 8, 2004. Accepted August 26, 2004.

The cause of latitudinal gradients in species diversity remains controversial (e.g., Rohde 1992, 1997). After reviewing several previously proposed hypotheses, Rohde (1992) concluded that this gradient most likely results from the increasing solar energy available in lower latitudes. He argued that (1) the increase in solar radiation leads to higher mutation rates, and (2) the increase in temperature results in shorter generation times, possibly leading to a higher mutation rate and/or faster selection. In combination, these factors lead to a faster rate of evolution for equatorial taxa relative to taxa inhabiting higher latitudes, which in turn leads to greater rates of speciation and results in the observed gradients (termed the climate-speciation hypothesis; Bromham and Cardillo 2003). This hypothesis predicts an indirect relationship between rate of molecular evolution and latitude or a correlate of latitude such as mean annual temperature (MAT; see Currie 1991).

To the best of our knowledge, no explicit tests of Rohde's hypothesis (that latitude should exert an influence on rate of molecular evolution) existed until recently (Bromham and Cardillo 2003; Wright et al. 2003). In their study of 45 phylogenetically independent pairs of bird species, Bromham and Cardillo (2003) found no evidence for a correlation between latitude and rate of molecular evolution. Conversely, in an analysis of 23 Pacific species of the plant subgenus *Mearnsia*, Wright et al. (2003) argued for a correlation between rate of molecular evolution and the MAT of the species' range.

Pacific species of *Mearnsia* occur on New Guinea (NG), the Philippines (Ph), the Solomon Islands (SI), New Zealand (NZ), and New Caledonia (NC). Wright et al. (2003) generated a phylogeny for the 23 Pacific species and determined the MAT of the midpoint of the elevational and latitudinal range of each species (Fig. 1). The average MAT of members of the NG/Ph/SI (equatorial) radiation was 7.1°C higher than the average MAT of the higher latitude NZ/NC (subequatorial) species. Wright et al. (2003) also noted that there are an increased number of substitutions per site (measured as branch length on a maximum likelihood phylogeny) between the terminal taxa of the equatorial radiation and the ancestor of all Pacific species (mean = 0.0465, $n = 6$) relative to this same measure for the subequatorial species (mean = 0.0172, $n = 17$). Because this method calculates branch lengths of each species by summing the number of substitutions per site from the common ancestor of all Pacific species to each terminal taxon, each species is treated as an independent unit despite the shared ancestry of the clade members. For example, the branch subtending the equatorial clade is included in the branch length tabulation of all six members of this clade (Fig. 1, branch A). To compensate for the lack of independence, Wright et al. (2003) conducted a one-tailed *t*-test using only one degree of freedom. Finding a significant result ($0.025 < P < 0.05$), the authors concluded that there was a correlation between rate of molecular evolution and MAT. However, we question the appropriateness of simply reducing the number of degrees of freedom in a *t*-test to address this problem. In the situation in which one or a few mutations occurred in the ancestor of all equatorial *Mearnsia* species that elevated the mutation rate, one would simply be comparing the rate of molecular evolution in those species with this mutation to those without—an effect in no way caused by the environment. Therefore, we explored alternative methods of analysis using phylogenetically independent contrasts to look for an effect of MAT on rate of molecular evolution across this entire subgenus.

Re-analysis of Mearnsia Dataset

After visual inspection of the ITS and ETS alignment of Wright et al. (2003), we noted several regions where sequences of one to a few individuals were shifted by a few bases. Therefore, we began by realigning the sequences using

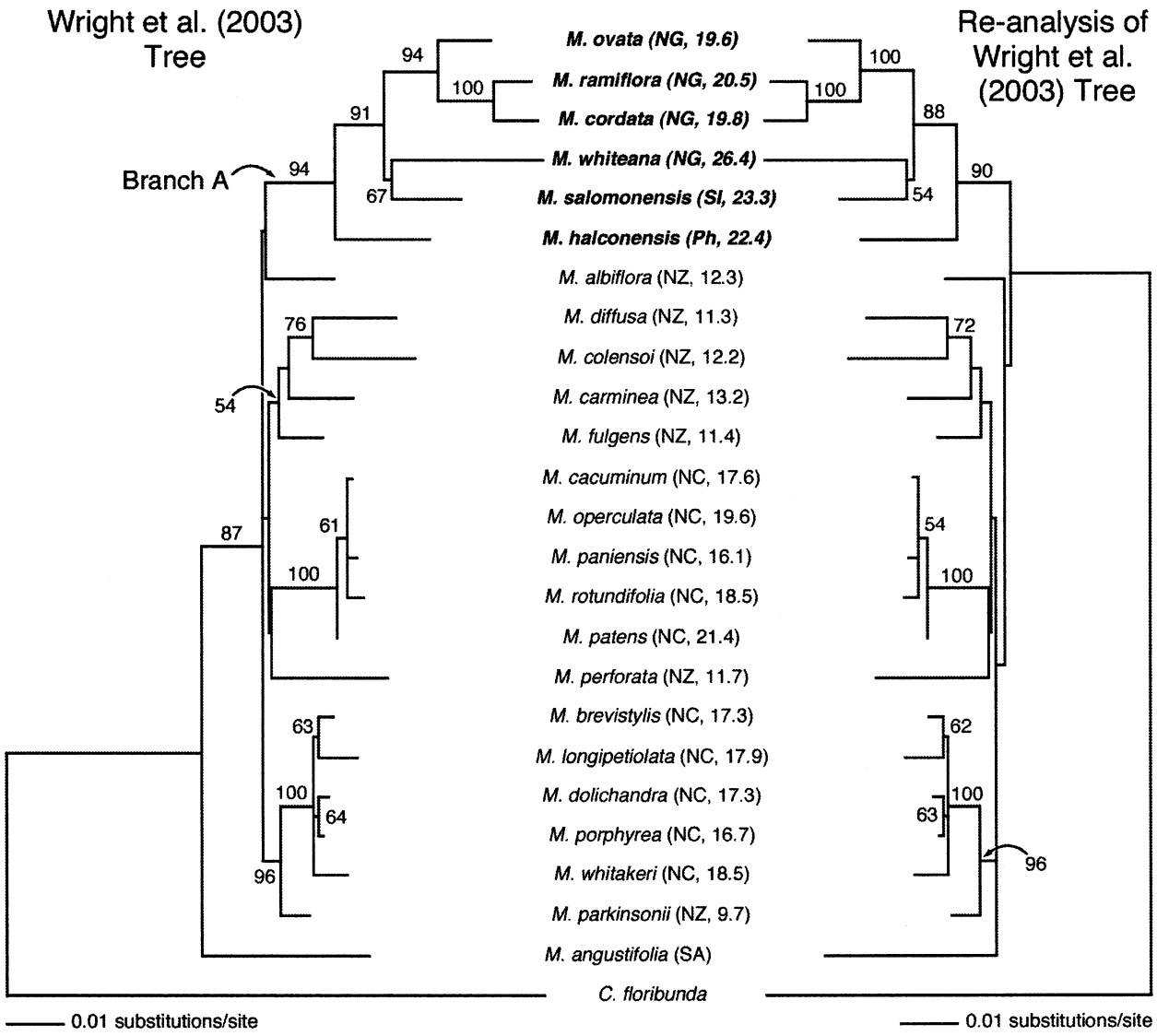

FIG. 1. Maximum likelihood phylogenies of *Metrosideros* subgenus *Mearnsia* species from the alignment used by Wright et al. (2003) (left tree) and from a realignment of their data (right tree). The scales are the same for both trees. Equatorial species are in bold. Species are followed by locality abbreviation (see text) and mean annual temperature (°C). Note the different reconstructions of *M. angustifolia*.

MacClade version 4.05 (Maddison and Maddison 2002). Changes were primarily in the alignment of *Metrosideros angustifolia* and *Cloezia floribunda* sequences relative to the remaining species. The realigned dataset included 111 parsimony-informative characters, whereas the alignment of Wright et al. (2003) contained 120 parsimony-informative characters. Phylogenetic analyses were conducted in PAUP* version 4.0b10 (Swofford 1999). We found the most-parsimonious trees using a branch-and-bound algorithm. Analysis of the realigned dataset yielded four most-parsimonious trees of 412 steps (CI = 0.745, RI = 0.708) whereas Wright et al.'s (2003) alignment yielded 130 most-parsimonious trees of 446 steps (CI = 0.762, RI = 0.714). Nonparametric bootstrapping was used to assess branch support by generating 1000 pseudoreplicate datasets and analyzing each replicate using a heuristic algorithm with a stepwise addition starting tree (generated through ''as-is'' sequence addition) and tree-bisection-reconnection (TBR) branch swapping. We then tested various models of evolution using Modeltest version 3.06 (Posada and Crandall 1998) under the Akaike information criterion (AIC) on one of the four most-parsimonious trees. The HKY + Γ model was chosen as the most appropriate, and the relevant parameters were optimized on the same most-parsimonious tree. These parameter values were then fixed and used to search among maximum likelihood (ML) trees with TBR branch swapping, using the four most-parsimonious trees as starting trees. The ML tree resulting from this search was then used to reoptimize parameter values and begin another search for a better ML tree. This process of successive approximation was continued until two identical ML trees were returned. The ML tree resulting from this search on the adjusted alignment was largely congruent with the tree given by Wright et al. (2003) aside from the placement of *M. angustifolia* (Fig. 1).

TABLE 1. Taxa used for maximum likelihood branch length and relative rates comparisons (*Metrosideros*). Outgroups were used for relative rates comparisons based on genetic distances.

Comparison	Outgroup	Taxon 1	Taxon 2
1	*M. carminea*	*M. colensoi*	*M. diffusa*
2	*M. whitakeri*	*M. longipetiolata*	*M. brevistylis*
3	*M. whitakeri*	*M. dolichandra*	*M. porphyrea*
4	*M. ovata*	*M. ramiflora*	*M. cordata*
5	*M. ovata*	*M. whiteana*	*M. salomonensis*
6	*M. patens*	*M. carminea*	*M. fulgens*
7	*M. fulgens*	*M. whitakeri*	*M. parkinsonii*
8	*M. parkinsonii*	*M. halconensis*	*M. ovata*
9	*M. fulgens*	*M. patens*	*M. perforata*

Independent Contrasts Analyses

For the following comparative analyses, the outgroup *C. floribunda* was removed because we were primarily interested in rate variation within the ingroup. Additionally, for most purposes (aside from rate smoothing; see below) *M. angustifolia* was removed because MAT data are not available for it. We first analyzed rate variation by calculating independent contrasts of both rate of molecular evolution and MAT between terminal taxa (similar to Bromham and Cardillo 2003). Contrasts were always calculated as the trait value of the higher MAT species minus the trait value of the lower MAT species. Therefore, all contrasts in MAT were positive, whereas contrasts in rate of substitution were only positive if the species with the higher MAT also had a higher rate. Rate contrasts were calculated both as the difference in terminal branch lengths on the maximum likelihood tree and as the difference in genetic distance from an outgroup (HKY + Γ corrected) chosen from the ML tree (relative rates test; Sarich and Wilson 1973). Species pairs and outgroups are given in Table 1. To avoid any potential node density effects, species pairs were pruned from the tree once contrasts were calculated between monophyletic sister species, and a new ML tree was found using the successive approximation procedure described previously. This pruning process continued until the maximum number of pairs had been examined. The species *Metrosideros cacuminum*, *M. operculata*, *M. paniensis*, and *M. rotundifolia* were not used for these tests because of the ambiguous relationships among them (Fig. 1). Sign tests and Wilcoxon signed-ranks tests were used to test for any deviations from a mean of zero among these contrasts. These tests did not reject a mean of zero (sign test, $P = 0.254$; Wilcoxon signed-rank test, $P = 0.50$; similar results were obtained when using the original alignment of Wright et al. 2003), which suggests that there is no effect of MAT on rate of molecular evolution.

We also looked for significant effects of MAT on rate variation using the independent contrasts approach of Felsenstein (1985). This approach involves the reconstruction of ancestral node trait values, which requires a measure of absolute rate of evolution for the terminal species that is independent of time (both branch length on a ML tree and genetic distances are measures of the rate of molecular evolution multiplied by time). To tease apart these factors we generated an ultrametric tree, based on our maximum likelihood tree, using the penalized likelihood rate smoothing approach of Sanderson (2002) as implemented in the program r8s version 1.06 (Sanderson 2003). A smoothing parameter of one was chosen as optimal for our ingroup using the cross-validation option (the Powell algorithm was used for optimization, and smoothing parameters ranging from 1 to 10,000 were tested). Cross-validation prunes off terminal branches one at a time, estimates their length based on rate smoothing across the remaining taxa, and then calculates the sum of errors between the actual and predicted branch lengths. Our chosen smoothing parameter minimized these errors. The most optimal solution for rate smoothing across the ingroup from 20 independent starting points was used (the stability of each optimization was checked with 1000 perturbations of parameter values by a factor of 0.001). The age of the ingroup root was fixed at 1.00 and all other nodes were assigned relative ages by optimization. For independent contrasts analysis, branch lengths on the ML tree topology were altered to reflect the difference in ages between nodes. The rates of molecular evolution assigned to each terminal branch through rate smoothing were then considered to be phenotypes of the corresponding terminal taxa. An independent contrast analysis across the ingroup (excluding *M. angustifolia*) between these rates and MAT values was then performed using the PDAP:PDTREE version 1.00 module (Midford et al. 2003) of Mesquite version 1.01 (Maddison and Maddison 2004). As suggested by Garland et al. (1992), the adequacy of the standardization procedure was checked by plotting the absolute value of each standardized independent contrast versus its standard deviation. No significant trend (except one extreme outlying point; see below) was noted, so branch lengths were not transformed.

Through least-squares regression of independent contrasts of rate of molecular evolution on contrasts of MAT, a marginally significant effect ($t = 1.76$, df = 21, one-tailed $P = 0.047$, $\beta = 3.7 \times 10^{-5}$) of MAT on rate of evolution was found, but this result was found to be entirely due to the position of one extreme outlier (Fig. 2). Once this contrast was removed from the analysis, the slope of the regression became negative and not significantly different from zero ($t = -0.62$, df = 20, one-tailed $P = 0.27$, $\beta = 6.1 \times 10^{-6}$). This outlying contrast corresponds to the root node of the ingroup and represents the contrast in rate of molecular evolution and MAT between the equatorial and subequatorial species of *Mearnsia*. Three other contrasts with similarly high MAT differences (i.e., >8.5) do not show a corresponding increase in rate of molecular evolution.

Even though Felsenstein's independent contrasts method is relatively robust to violations of assumptions regarding mode of evolution (Díaz-Uriarte and Garland 1996), we examined the possibility that biases in reconstructing ancestral node trait values caused spurious results in this situation given that MAT has an unknown pattern of change across species in this group. To do this, we examined only those contrasts between terminal taxa. The slope of this regression was also not significantly different than zero ($t = 0.02$, df = 8, one-tailed $P = 0.50$, $\beta = 2.3 \times 10^{-7}$).

By looking at a plot of rate of molecular evolution, in which rates are estimated through penalized likelihood rate smoothing (Sanderson 2002), as a function of MAT for each species without correcting for phylogeny (Fig. 3), it appears,

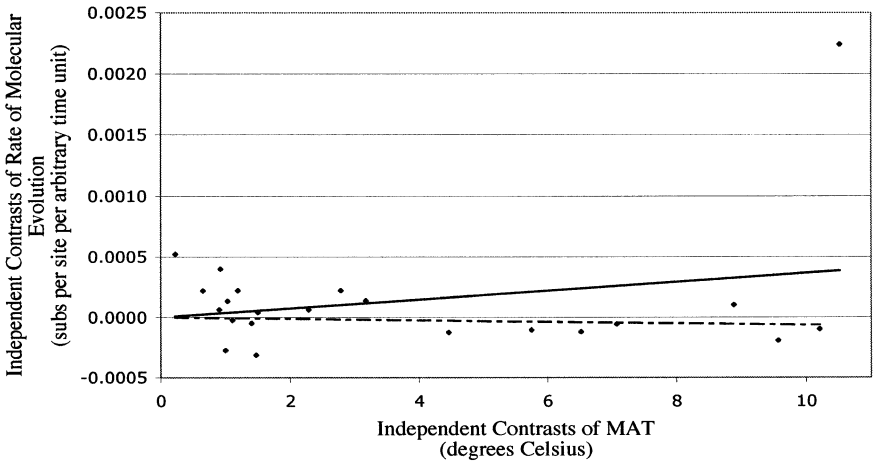

FIG. 2. The least-squares regression of contrasts in rate of molecular evolution as a function of contrasts in mean annual temperature (MAT; solid line). The regression is forced through the origin, as required by independent contrast techniques (Felsenstein 1985). The point in the upper right corner of the plot is clearly an outlier and the cause of the positive slope. Removal of this point results in a negative slope for the regression (dotted and dashed line). This point corresponds to the contrast between equatorial and subequatorial *Mearnsia* species.

as pointed out by Wright et al. (2003), that the rates of evolution for all equatorial species are higher than for any of the subequatorial species. This is true even for equatorial and subequatorial species occurring at nearly identical MATs. For instance, the subequatorial *M. patens* has a MAT higher than three (*M. ovata*, *M. ramiflora*, and *M. cordata*) of the six equatorial species, yet its rate of molecular evolution is drastically lower (Fig. 3). Additionally, there seems to be no trend in rate within either clade. The regression line in Figure 3 is, in effect, based only on the means of the two clades.

Conclusions

In light of the above analyses, an effect of phylogenetic nonindependence appears to be a plausible explanation of the greater substitution rate in the equatorial *Mearnsia*. However, given the small number of *Mearnsia* species examined, extracting a general pattern may be especially problematic. Mean annual temperature, or an ecological correlate of MAT, may impact rate of evolution, but determining the causal factor of rate variation is likely to require more data than is available in this case. If the correlation between MAT and rate of molecular evolution is relatively weak, then multiple contrasts with large differences may be necessary before a substantial change in rate of evolution is seen. Other environmental effects, such as mean annual precipitation, mean annual evapotranspiration, or the extent of annual variation in temperature, precipitation, and/or evapotranspiration (Currie and Paquin 1987; Currie 1991) may impact rates of evolution in different ways, providing confounding ''noise'' to estimates of ecological correlates of rate variation. Other en-

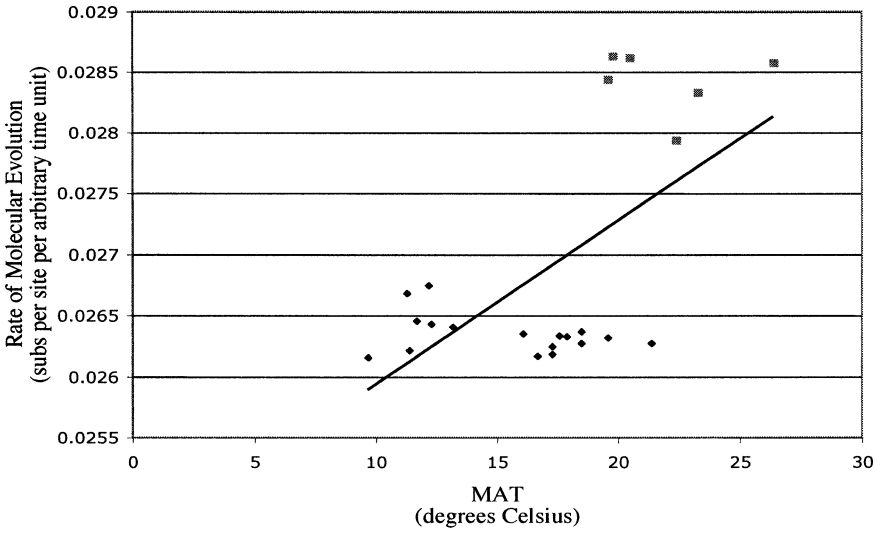

FIG. 3. The least-squares regression of the rate of evolution for extant equatorial (squares) and subequatorial (diamonds) *Mearnsia* species as a function of mean annual temperature (MAT) without correcting for phylogeny. Note that all equatorial species have a rate >0.0275.

vironmental variables may be the true cause of rate variation in Pacific species of *Mearnsia* but may be only weakly correlated with MAT. To address this issue, it would be ideal to simultaneously measure many ecological variables that could be examined using phylogenetically independent contrasts for correlations with rates of evolution.

Acknowledgments

We thank members of the Hill Country Student Journal Club at the University of Texas for stimulating discussions on this topic. D. Bickford, S. Pawar, and D. Hillis provided helpful comments on this manuscript, and D. Zwickl provided guidance with regards to analyses. We also thank S. Wright for providing us with the original aligned sequence data. JMB was supported by a Donald D. Harrington fellowship from the University of Texas. GBP was supported by a National Science Foundation IGERT fellowship in Computational Phylogenetics and Applications to Biology (DGE-0114387).

Literature Cited

Bromham, L., and M. Cardillo. 2003. Testing the link between the latitudinal gradient in species richness and rates of molecular evolution. J. Evol. Biol. 16:200–207.

Currie, D. J. 1991. Energy and the large-scale patterns of animal- and plant-species richness. Am. Nat. 137:27–49.

Currie, D. J., and V. Paquin. 1987. Large-scale patterns of species richness of trees. Nature 329:326–327.

Díaz-Uriarte, R., and T. Garland, Jr. 1996. Testing hypotheses of correlated evolution using phylogenetically independent contrasts: sensitivity to deviations from Brownian motion. Syst. Biol. 45:27–47.

Felsenstein, J. 1985. Phylogenies and the comparative method. Am. Nat. 125:1–15.

Garland, Jr., T., P. H. Harvey, and A. R. Ives. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. Syst. Biol. 41:18–32.

Maddison, D. R., and W. P. Maddison. 2002. MacClade 4. Ver. 4.05. Sinauer Associates, Sunderland, Ma.

Maddison, W. P., and D. R. Maddison. 2004. Mesquite: a modular system for evolutionary analysis. Ver. 1.01. Available via http://mesquiteproject.org.

Midford, P. E., T. Garland, Jr., and W. P. Maddison. 2003. PDAP package. Available via http://mesquiteproject.org/pdap_mesquite/.

Posada, D., and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818.

Rohde, K. 1992. Latitudinal gradients in species diversity: the search for the primary cause. Oikos 65:514–527.

———. 1997. The larger area of the tropics does not explain latitudinal gradients in species diversity. Oikos 79:169–172.

Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19:101–109.

———. 2003. r8s: analysis of rates of evolution (and other stuff). Ver. 1.06. Available via http://ginger.ucdavis.edu/r8s/.

Sarich, V. M., and A. C. Wilson. 1973. Generation time and genomic evolution in primates. Science 179:1144–1147.

Swofford, D. L. 1999. PAUP*. Phylogenetic analysis using parsimony (* and other methods). Ver. 4. Sinauer Associates, Sunderland, MA.

Wright, S. D., R. D. Gray, and R. C. Gardner. 2003. Energy and the rate of evolution: inferences from plant rDNA substitution rates in the western Pacific. Evolution 57:2893–2898.

Corresponding Editor: G. May